Remarks

Amendments to the Claims

Each of independent claims 1, 12, 33, 52, and 59 is amended to recite that the biological subject is "obtained" from a region of the mammal. Similarly, each of independent claims 9, 22, and 62 is amended to recite that the second sample is "obtained" from the patient. The specification supports these amendments on page 22, line 11 to page 23, line 2, where it describes obtaining biological subjects and samples.

Each of independent claims 9, 22, and 62 is amended to clarify that the recited steps are carried out at two different time points during a therapeutic regimen. The specification supports this amendment on page 56, lines 13-23, which describes taking samples from patients "at various time points during the course of the application of a treatment regimen."

New claims 65 and 66 recite that the mammal or patient, respectively, is a human, and new claims 67-82 recite human hepsin sequences SEQ ID NO:1 (nucleotide) and SEQ ID NO:2 (amino acid), which are disclosed in the specification as filed.

The amendments do not add new matter.

The Rejection of Claims 59-64 Under 35 U.S.C. § 112 ¶ 2

The Office Action maintains the rejection of claims 59-64 under 35 U.S.C. § 112 ¶ 2 because the recitations "first indirect measure" and "second indirect measure" allegedly are indefinite. Applicants respectfully traverse the rejection.

Under the second paragraph of 35 U.S.C. § 112, the relevant inquiry

> . . . is merely to determine whether the claims do, in fact, set out and circumscribe a particular area with a reasonable degree of precision and particularity. It is here where the definiteness of the

> language employed must be analyzed -- not in a vacuum, but always in light of the teachings of the prior art and of the particular application disclosure as it would be interpreted by one possessing the ordinary level of skill in the pertinent art.

In re Moore, 439 F.2d 1232, 1235, 58 C.C.P.A. 1042, 1046-47 (1971) (footnote omitted). The Office Action contends that "[t]he terms, 'first indirect measure' or 'second indirect measure' are not defined by the claim(s), the specification does not provide a standard for ascertaining the requisite degree, and one of ordinary skill in the art would not be reasonably apprised of the metes and bounds of such terms." Office Action at page 3. On the contrary, one of ordinary skill in the art would understand that, in the context of determining gene copy number, any method other than directly counting copies of genes is, by definition, an indirect measure of gene copy number. Indirect methods of detecting gene copy number were well known in the art when this application was filed. As the specification teaches, these methods include Southern blotting, *in situ* hybridization, comparative genomic hybridization (CGH), amplification-based assays (*e.g.*, a PCR-based TaqMan assay), and DNA microarray-based CGH (pages 41-43; Example 1). Each of these methods detects a measurement which reflects gene copy number and is, therefore, indirect. In the context of the specification and the knowledge of one of ordinary skill, the term "indirect measure" as used in claims 59-64 is definite.

Under 35 U.S.C. § 112, second paragraph, the claims must "reasonably apprise those skilled in the art both of the utilization and scope of the invention." *Georgia-Pacific Corp. v. United States Plywood Corp.*, 258 F.2d 124, 136, 118 U.S.P.Q. 122, 130 (2d Cir. 1958), *cert. denied*, 358 U.S. 884 (1958). Claims 59-64 serve this purpose. Claims 59-64 are therefore definite.

Applicants respectfully request withdrawal of the rejection.

Rejection of Claims 1-3, 9-12, 22-24, 33-35, and 39-64 Under 35 U.S.C. § 112 ¶ 1

Claims 1-3, 9-12, 22-24, 33-35, and 39-64 stand rejected under 35 U.S.C. § 112 ¶ 1 as insufficiently described. The Office Action asserts the term "hepsin" is broader than the description supports. Applicants respectfully traverse the rejection.

What is required to satisfy the written description requirement depends on the nature of the invention claimed. *In re DiLeone*, 436 F.2d 1404, 1405, 168 U.S.P.Q. 592, 593 (C.C.P.A. 1971). Independent claims 1, 52, and 59 recite a hepsin gene copy number in a biological subject obtained from a region of a mammal that is suspected to be precancerous or cancerous. Independent claims 9 and 62 recite a hepsin gene copy number in a sample of precancerous of cancer cells obtained from a patient. Independent claim 12 recites a test level of hepsin mRNA expression in a biological subject obtained from a region of the mammal that is suspected to be precancerous or cancerous. Independent claim 22 recites expression levels of at least one of hepsin mRNA or hepsin protein in samples of a biological subject comprising precancerous or cancer cells obtained from a patient. Independent claim 33 recites detecting a test hepsin protein expression level in a biological subject obtained from a region of the mammal that is suspected to be precancerous or cancerous. In each case, the recited hepsin gene, protein, or mRNA is one which naturally occurs in the mammal.

The naturally occurring mammalian hepsin protein and gene were well known in the art at the February 14, 2001 priority date of this application. *See, e.g.*, Kawamura[1] and Vu[2] (mouse

[1] Kawamura *et al.*, "Complete nucleotide sequence, origin of isoform and functional characterization of the mouse hepsin gene," Eur J Biochem. 1999 Jun;262(3):755-64.

[2] Vu *et al.*, "Identification and Cloning of the Membrane-associated Serine Protease, Hepsin, from Mouse Preimplantation Embryos," *J. Biol. Chem. 272*, 31315-20, 1997.

hepsin); Kazama[3] and Leytus[4] (human hepsin); and Farley (rat hepsin),[5] each provided with the accompanying Information Disclosure Statement. An adequate written description of a gene which is well known in the art does not require a structural recitation either in the specification or in the claims. *See Capon v. Eshhar*, 418 F.3d 1349, 1360-61, 76 U.S.P.Q.2d 1078, 1087 (Fed. Cir. 2005) ("the Board erred in ruling that § 112 imposes a *per se* rule requiring recitation in the specification of the nucleotide sequence of claimed DNA, when that sequence is already known in the field."). Thus, the fact that the claims do not recite a sequence identifier "relating to hepsin" as noted in the Office Action at page 4 ¶ 1 does not mean that the claims lack written description. Under *Capon*, a sequence identifier is not required to describe the term "hepsin."

The Office Action cites nine different GenBank Accession Numbers as disclosing hepsin genes. Page 4 ¶ 1. The hepsin coding sequences with these Accession Numbers, however, are not "highly variant," as the Office Action contends. *Id.*

Accession Numbers NM_017112 and X70900 are rat hepsin coding sequences; a BLAST alignment of these sequences shows that the two sequences are 100% identical (Exhibit 1). Accession Numbers NM_008281 and AF030065 are mouse hepsin coding sequences; a BLAST alignment of these sequences shows that the two sequences are 100% identical (Exhibit 2). The mouse and rat sequences are 94% identical (Exhibit 3).

Accession Numbers BC025716.1, NM_182983.1, NM_002151.1, X07732, and X07002.1 are human hepsin coding sequences. BLAST alignments of SEQ ID NO:1 with each of the other

[3] Kazama *et al.*, "Hepsin, a putative membrane-associated serine protease, activates human factor VII and initiates a pathway of blood coagulation on the cell surface leading to thrombin formation," J Biol Chem. 1995 Jan 6;270(1):66-72.

[4] Leytus *et al.*, "A novel trypsin-like serine protease (hepsin) with a putative transmembrane domain expressed by human liver and hepatoma cells," Biochemistry 27, 1067-1074, 1988.

[5] Farley *et al.*, "Cloning and sequence analysis of rat hepsin, a cell surface serine proteinase," Biochim Biophys Acta. 1993 Jun 25;1173(3):350-2.

human hepsin coding sequences are provided in Exhibits 4-8. SEQ ID NO:1 is 100% identical over 1593 contiguous nucleotides and 100% identical over 193 contiguous nucleotides with each of splice variants NM_182983.1 (Exhibit 5) and X07732 (Exhibit 7). SEQ ID NO:1 is 99% identical with BC025716.1 (Exhibit 4), 100% identical with NM_002151.1 (Exhibit 6), and 100% identical with 1199 contiguous nucleotides of X07002.1 (Exhibit 8). SEQ ID NO:1 is 83% identical to a long stretch (1318 nucleotides) of the rat and mouse hepsin coding sequences (Exhibit 9).

Mammalian hepsin coding sequences were known in the art when this application was filed, and none of the sequences cited in the Office Action are "highly variant." The specification therefore satisfies the written description requirement with respect to the term "hepsin" as recited in the rejected claims.

Applicants respectfully request withdrawal of the rejection.

Enablement Rejections

All of the pending claims stand rejected under 35 U.S.C. § 112 ¶ 1 as not enabled for their full scope. In each case, the Office Action contends that the specification enables the claimed methods for *ex vivo* use but not for use *in vivo*. Office Action at page 7 ¶ 1. Applicants respectfully traverse the rejection.

To advance prosecution, each of the independent claims has been amended to clarify that the recited method steps are carried out *ex vivo*. Each of the independent diagnostic method claims (1, 12, 33, 52, and 59) now recites that the biological subject is "obtained from" a region of the mammal suspected to be precancerous or cancerous. Similarly, each of the independent methods of monitoring treatment efficacy (9, 22, and 62) recites that the sample is "obtained

from" a patient. Because the Office Action acknowledges this subject matter is enabled, Applicants respectfully request withdrawal of the rejection.

Respectfully submitted,

BANNER & WITCOFF, LTD.

/Lisa M. Hemmendinger/

Date: October 27, 2006 By: _____

Lisa M. Hemmendinger
Registration No. 42,653

Customer No. 22907

Exhibit 1. BLAST alignment of NM_017112 and X70900

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 30 expect: 10.0000 wordsize: 11 Filter: ☐ View option: Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
☐ Show CDS translation Align

Sequence 1: gi|8393559|ref|NM_017112.1|Rattus norvegicus hepsin (Hpn), mRNA
Length = 1739 (1 .. 1739)

Sequence 2: gi|57928|emb|X70900.1|RNHEPAR Rattus norvegicus mRNA for hepsin
Length = 1739 (1 .. 1739)

NOTE: Bitscore and expect value are calculated based on the size of the nr database.

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

```
 Score = 3944 bits (1739),  Expect = 0.0
 Identities = 1739/1739 (100%), Gaps = 0/1739 (0%)
 Strand=Plus/Plus

Query  1     GCAGGCCCACGCTGCTGGCTGCTGCTGCCACCCTTGCCTCCGGGCTGCCCGCTGCTGT  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     GCAGGCCCACGCTGCTGGCTGCTGCTGCCACCCTTGCCTCCGGGCTGCCCGCTGCTGT  60

Query  61    GGGGACACCATGCCCCTTCCCAGGCCCGGACTAACCCCAAACCTGCACCAGTTCCGGC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    GGGGACACCATGCCCCTTCCCAGGCCCGGACTAACCCCAAACCTGCACCATTCCGGC  120

Query  121   GAACCCCAGGGTTCCGCCCCAGCCCAAACGGTCAACCTGGGAATCATTAACAAGAGTCCC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  121  GAACCCCAGGGTTCCGGCCCAGCCCAACAGGTCAACCTGGGAATCATTAACAAGAGTCCC  180

Query  181  TGACATGGCGAAGGAGGGTGGCGGACTGCACCATGCTGTTCCAGACCCAAGGTGGCAGC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  TGACATGGCGAAGGAGGGTGGCCGGACTGCACCATGCTGTTCCAGACCCAAGGTGGCAGC  240

Query  241  TCTCACTGTGGGGACCCTGCTGTTCCTGACAGGCATTGGGGCTGCGTCCTGGGCCATTGT  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241  TCTCACTGTGGGGACCCTGCTGTTCCTGACAGGCATTGGGGCTGCGTCCTGGGCCATTGT  300

Query  301  GACCATCCTACTACGGAGTGACCAGGAGCCACTGTACCAAGTGCAGCTCAGTCCTGGGGA  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301  GACCATCCTACTACGGAGTGACCAGGAGCCACTGTACCAAGTGCAGCTCAGTCCTGGGGA  360

Query  361  CTCTCGACTTTTGGTGTTGGACAAGACAGAGGGAACGTGGAGGCTCCTGTGCTCCTCAGG  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361  CTCTCGACTTTTGGTGTTGGACAAGACAGAGGGAACGTGGAGGCTCCTGTGCTCCTCAGG  420

Query  421  GTCCAAGGCCAGGGTAGCAGGGCTCGGCTGTGAGGAGATGGGCTTCTCCAGGGGTCTGGC  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421  CTCCAAGGCCAGGGTAGCAGGGCTCGGCTGTGAGGAGATGGGCTTCTCCAGGGGTCTGGC  480

Query  481  GCACTCAGAGGTGGAGTGCGAACCGCGGGCGCCAACGGCACATCGGGCTTCTTCTGCGT  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481  GCACTCAGAGGTGGAGTGCGAACCGCGGGCGCCAACGGCACATCGGGCTTCTTCTGCGT  540

Query  541  GGACGAGGGCGGTCTGCCTCTGGCTCAGCGGTTGCTGGATGTCATCTCTGTATGCGACTG  600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541  GGACGAGGGCGGTCTGCCTCTGGCTCAGCGGTTGCTGGATGTCATCTCTGTATGCGACTG  600

Query  601  TCCTAGAGGCCGATTCCTGACTGCCACCTGCCAAGACTGTGGCCGCCAGGAAGCTGCCGGT  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601  TCCTAGAGGCCGATTCCTGACTGCCACCTGCCAAGACTGTGGCCGCCAGGAAGCTGCCGGT  660

Query  661  GGATCGCATTGTGGGGGGCCAGGACAGCAGCCTGGGAAGATGGCCATGGCAGGTCAGCCT  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661  GGATCGCATTGTGGGGGGCCAGGACAGCAGCCTGGGAAGATGGCCATGGCAGGTCAGCCT  720

Query  721  GGGTTATGATGGGACCCACCTCTGTGGGGGATCCCCTGCTGTCCGGGGACTGGGTACTGAC  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721  GGGTTATGATGGGACCCACCTCTGTGGGGGATCCCCTGCTGTCCGGGGACTGGGTACTGAC  780

Query  781  CGCTGCACACTGCTTTCCAGAGAGGAACCGGGTCCTGTCTCGGTGGCGAGTATTTGCTGG  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  781  CGCTGCACACTGCTTTCCAGAGAGGAACCGGGTCCTGTCTCGGTGGCGAGTATTTGCTGG  840

Query  841  TGCTGTAGCCCGGACCTCACCTCATGCGTGCAGCTGGGGGTTCAGGCTGTGATCTATCA  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  841  TGCTGTAGCCCGGACCTCACCTCATGCGTGCAGCTGGGGGTTCAGGCTGTGATCTATCA  900

Query  901  TGGGGGCTACCTTCCCTTTCGAGACGCTACTATCGACGAAAACAGCAATGACATTGCCCT  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  901  TGGGGGCTACCTTCCCTTTCGAGACGCTACTATCGACGAAAACAGCAATGACATTGCCCT  960

Query  961  GGTCCACCTCTCTAGTCCCTGCCTCTCACAGAATACATCAGCCGGGTTTGTCTCCTGC  1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  961  GGTCCACCTCTCTAGCTCCTGCCTCTCACAGAATACATCAGCCGGGTTTGTCTCCTGC  1020

Query  1021  TGTGGGGACAGGCCCTGGTGGACGGCAAGGTCTGTACAGTGACCGGCTGGGGTAACACACA  1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  1021  TGCGGGACAGGCCCTGGTGGACGGCAAGGTCTGTACAGTGACGGGTGGGGTAACACACA  1080

Query  1081  GTTCTATGGCCAGCAAGCTGTGGTGCTCCAAGAGGCCCGGGTCCCCATCATAAGCAACGA  1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1081  GTTCTATGGCCAGCAAGTTTTGGTGCTCCAAGAGGCCCGGGTCCCCATCATAAGCAACGA  1140

Query  1141  AGTTTGCAACAGCCCCGACTTCTACGGGAATCAGATCAAACCCAAGATGTTCTGTGGTGG  1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1141  AGTTTGCAACAGCCCCGACTTCTACGGGAATCAGATCAAACCCAAGATGTTCTGTGGTGG  1200

Query  1201  CTATCCTGAGGGTGGTATTGATGCATGCCAGGGTGACAGCGGAGGCCACTTTGTATGTGA  1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201  CTATCCTGAGGGTGGTATTGATGCATGCCAGGGTGACAGCGGAGGCCACTTTGTATGTGA  1260

Query  1261  GGACAGAATCTCTGGAACATCAAGATGGCGGCTCTGCGGCATTGTAAGCTGGGGTACGGG  1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1261  GGACAGAATCTCTGGAACATCAAGATGGCGGCTCTGCGGCATTGTAAGCTGGGGTACGGG  1320

Query  1321  CTGTGTTTTGGCCCGGAAGCCGGGAGTGTACACCAAAGTCATTGACTTCCGGGAGTGGAT  1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1321  CTGTGTTTTGGCCCGGAAGCCGGGAGTGTACACCAAAGTCATTGACTTCCGGGAGTGGAT  1380

Query  1381  CTTCCAGGCGATAAAGAGTGACTCCGAAGCTACCGGCATGGTAACTCAGCGCTGACGCCG  1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1381  CTTCCAGGCGATAAAGACTCACTCCGAAGCTACCGGCATGGTAACTCAGCGCTGACGCCG  1440

Query  1441  CCTCATCGGCTGCTCCGCGCTGCTCCAGCATCCAGAGTCAGAGTTGGTCTGGTGGCTCCA  1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1441  CCTCATCGGCTGCTCCGCGCTGCTCCAGCATCCAGAGTCAGAGTTGGTCTGGTGGCTCCA  1500

Query  1501  GCGGACGTGGCAGGCTCCACACTGGGCTTCACATGGAACGGTTTTCTGCTCGGATCCAG  1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1501  GCGGACGTGGCAGGCTCCACACTGGGCTTCACATGGAACGGTTTTCTGCTCGGATCCAG  1560

Query  1561  TCCATAGATCCAAGGATGCTGGGTCCAAGGACCTCTCTCCACAGTGGCCGGCCACTCA  1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1561  TCCATAGATCCAAGGATGCTGGGTCCAAGGACCTCTCTCCACAGTGGCCGGCCACTCA  1620

Query  1621  ATCCCAGGGCCATTGGGTCACCCTCCCACCCCAATGTAAATATTACTCTGTCCTCTGGGG  1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1621  ATCCCAGGGCCATTGGGTCACCCTCCCACCCCAATGTAAATATTACTCTGTCCTCTGGGG  1680

Query  1681  GGTGCTTTCGAGGGCGCCCTTGTGCGGATGCTCTTTAAATATTAAAGGTGGTTTTGATT  1739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1681  GGTGCTTTCGAGGGCGCCCTTGTGCGGATGCTCTTTAAATATTAAAGGTGGTTTTGATT  1739
```

```
CPU time:     0.02 user secs.       0.00 sys. secs          0.02 total secs.

Lambda      K      H
    1.33   0.621     1.12

Gapped
Lambda      K      H
    1.33   0.621     1.12


Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 5, Extension: 2
```

```
Number of Sequences: 1
Number of Hits to DB: 482
Number of extensions: 1
Number of successful extensions: 1
Number of sequences better than 10.0: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Length of query: 1788
Length of database: 17,904,826,723
Length adjustment: 27
Effective length of query: 1712
Effective length of database: 17,904,826,703
Effective search space: 30656804313824
Effective search space used: 30656804313824
X1: 11 (21.1 bits)
X2: 15 (28.8 bits)
X3: 26 (50.0 bits)
S1: 14 (27.6 bits)
S2: 22 (43.6 bits)
```

Exhibit 2.

NCBI **Blast 2 Sequences results**

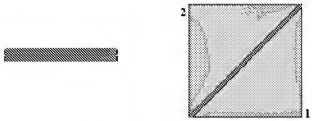
PubMed Entrez BLAST Genome Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 10.0000 wordsize: 11 Filter: ☐ View option: Standard
Masking character option: X for protein, N for nucleotide Masking color option: Black
☐ Show CDS translation Align

Sequence 1: gi|2605036|gb|AF030065.1|AF030065
Length = 1781 (1 .. 1781)

Sequence 2: gi|6680266|ref|NM_008281.1|Mus musculus hepsin (Hpn), mRNA
Length = 1781 (1 .. 1781)



NOTE:Bitscore and expect value are calculated based on the size of the nr database.

NOTE:If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

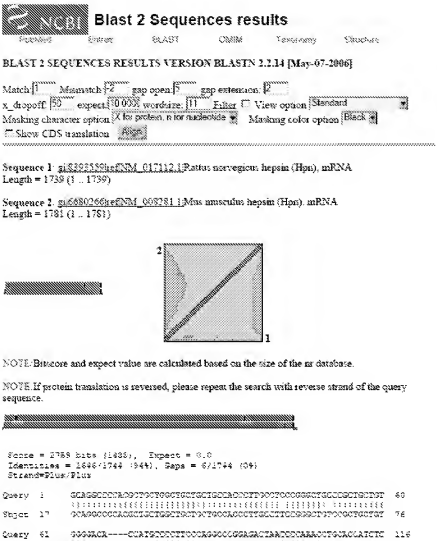

```
 Score = 3424 bits (1781),  Expect = 0.0
 Identities = 1781/1781 (100%), Gaps = 0/1781 (0%)
 Strand=Plus/Plus

Query  1     CGGAGGAGGGGGCTCGGCAGGCCCCACGCTGCTGGCTGCTGCTGCCACCCTTGCCTTCCG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     CGGAGGAGGGGGCTCGGCAGGCCCCACGCTGCTGGCTGCTGCTGCCACCCTTGCCTTCCG  60

Query  61    GGCTGTCGCTGCTGTGGGGACAGACACCATGCCCCTGCCCAGGCCCGGAGACTAACCCC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    GGCTGTCCGCTGCTGTGGGGACAGACACCATGCCCCTGCCCAGGCCCGGAGACTAACCCC  120

Query  121   AAACCTGCACCATGTCCGGCGAACCCCAGGGTTCCGCCCCAGCCCAACAGGTCAACCTGG  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  121  AAACCTGCACCATCTCCGGCGAAACCCAGGGTTCGCCCCAGCCCAACAGGTCAACCTGG  180

Query  181  GAATCATTAACAAGAGTCCCTGACATGGCGAAGGAGGGTGGCCGGACTGCAGCATGCTGC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  GAATCATTAACAAGAGTCCCTGACATGGCGAAGGAGGGTGGCCGGACTGCAGCATGCTGC  240

Query  241  TCCAGACCCAAGGTGGCAGCTCTCATTGTGGGTACCCTGCTGTTCCTGACAGGCATTGGG  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241  TCCAGACCCAAGGTGGCAGCTCTCATTGTGGGTACCCTGCTGTTCCTGACAGGCATTGGG  300

Query  301  GCGGCGTCCTGGGCCATTGTGACCATCCTACTGCAGAGTGAGCAGCGACCACTGTACCAA  360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301  GCGGCGTCCTGGGCCATTGTGACCATCCTACTGCAGAGTGAGCAGCGACCACTGTACCAA  360

Query  361  GTGCAGCTCAGTCAGGGGACTCACGAGTTGCAGTGTTGGACAAGACGGAGGAGGTACGTGG  420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361  GTGCAGCTCAGTCCAGGGGACTCACGACTTGCAGTGTTGGACAAGACGGAGGAGGTACGTGG  420

Query  421  AGGCTACTGTGGTCCTCACGGTCCAATGCCAGGGTGGCAGGCCTGGCTGTGAGGAGATG  480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421  AGGCTACTGTGGTCCTCACGGTCCAATGCCAGGGTGGCAGGCCTGGCTGTGAGGAGATG  480

Query  481  GGCTTTCTCAGGGGTCTGGCGCACTCGGAGTGGATGTGCGCACTGCGGGCGCCAACGGC  540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481  GGCTTTCTCAGGGGTCTGGCGCACTCGGAGCTGGATGTGCGCACTGCGGGCGCCAACGGC  540

Query  541  ACATCGGGCTICTTTTGCGTGGACGAGGGCGGACTGCCTCTGGCTCAGAGGTTGCTGGAI  600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  541  ACATCGGGCTICTTTTGCGTGGACGAGGGCGGACTGCCTCTGGCTCAGAGGTTGCTGGAT  600

Query  601  GTCATCTCTGTATGTGAGTGTCCTAGAGGCCGATTCCTGACTGCCACCTGCCAAGACTGT  660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  601  GTCATCTCTGTATGTGAGTGTCCTAGAGGCCGATTCCTGACTGCCACCTGCCAAGACTGT  660

Query  661  GGCGCAGGAAGCTGCCGGTGGACGCATTGTGGGGGGCCAGGACAGCARTCTGGGAAGG  720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  661  GGCGCAGGAAGCTGCCGGTGGACGCATTGTGGGGGGCCAGGACAGCAGTCTGGGAAGG  720

Query  721  TGGCCGTGGCAGGTCAGCCTGCGTIATGATGGGACCCACCTCTGTGGGGGGTCCCTGCTG  780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  721  TGGCCGTGGCAGGTCAGCCTGCGTIATGATGGGACCCACCTCTGTGGGGGGTCCCTGCTG  780

Query  781  TCTGGGGACTGGGTGCTGACTGCTGCACATTGCTTTCCAGAGCGGAACCGGGTCCTGTCT  840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  781  TCTGGGGACTGGGTGCTGACTGCTGCACATTGCTTTCCAGAGCGGAACCGGGTCCTGTCT  840

Query  841  CGGTGGCGAGTATTTGCTGGTGCTGTAGCCCGGACCTCACCCCATGCTGTGCAACTGGGG  900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  841  CGGTGGCGAGTATTTGCTGGTGCTGTAGCCCGGACCTCACCCCATGCTGTGCAACTGGGG  900

Query  901  GTTCAGGTGTGATCTATCATGGGGGCTACCTTCCTTTCGAGACCCTACTATTGACGAA  960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  901  GTTCAGGTGTGATCTATCATGGGGGCTACCTTCCTTTCGAGACCCTACTATTGACGAA  960

Query  961  AACAGCAATGACATTGCCTTGGTCCACCTCTCTAGCTCCCTGCCTCTCACAGAATACATC  1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  961  AACAGCAATGACATTGCCTTGGTCCACCTCTCTAGCTCCCTGCCTCTCACAGAATACATC  1020

Query  1021  CAGCCAGTGTGTCTCCCTGCTGCGGGACAGGCCCTGGTGGATGCCAAGGTCTGTACTGTG  1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  1021  CAGCCAGTGTGTCTCCCTGCTGCGGGACAGGCCCTGGTGGATGGTAAGGTCTGTACTGTG  1080

Query  1081  ACCGGCCTGGGGTAACACACAAGTTCTATGGCCAACAGGCTATGGTGCTCCAAGAGGCCGG  1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1081  ACCGGCCTGGGGTAACACACAAGTTCTATGGCCAACAGGCTATGGTGCTCCAAGAGGCCGG  1140

Query  1141  GTTCCCATCATAAGCAACGAAGTTTGCAACAGGCCCGGATTCTACGGGAATCAGATCAAG  1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1141  GTTCCCATCATAAGCAACGAAGTTTGCAACAGGCCCGGATTCTACGGGAATCAGATCAAG  1200

Query  1201  CCCAAGATGTCTGTGCTGGCTATCCTGAGGGTGGCATTGATGCGTGCCAGTGCGACAGC  1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201  CCCAAGATGTCTGTGCTGGCTATCCTGAGGGTGGCATTGATGCGTGCCAGTGCGACAGC  1260

Query  1261  GGAGGCCCCTTTGTGTGTGAAGACAGCATCTCTGGGACATCAAGGTGGCGGCTATGTGGC  1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1261  GGAGGCCCCTTTGTGTGTGAAGACAGCATCTCTGGGACATCAAGGTGGCGGCTATGTGGC  1320

Query  1321  ATTGTAAGCTGGGGTACCGGCTGTGCTTTGGCCCGGAAGCCAGGAGTGTACACCAAAGTC  1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1321  ATTGTAAGCTGGGGTACCGGCTGTGCTTTGGCCCGGAAGCCAGGAGTGTACACCAAAGTC  1380

Query  1381  ACTGACTTCCGGGAGTGGATCTTCAAGGCCATAAAGACTCACICCGAAGCCAGTGGCATG  1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1381  ACTGACTTCCGGGAGTGGATCTTCAAGGCCATAAAGACTCACTCCGAAGCCAGTGGCATG  1440

Query  1441  GTAACTCAGCCTGATCCCGCCATCTCGCTGCTCGGTGTGTGCACTAGCATCCAGAGTC  1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1441  GTGACTCAGCCTGATCCCGCCATCTCGCTGCTCGGTGTGTGCACTAGCATCCAGAGTC  1500

Query  1501  AGAGTTGGTCTGGTGGCTCCAGCCCCAGTGGTAGGCTCCACACTGGGCCTCACATGGAA  1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1501  AGAGTTGGTCTGGTGGCTCCAGCCCCAGTGGTAGGCTCCACACTGGGCCTCACATGGAA  1560

Query  1561  TGGTTTCCTGCTCAGATCCAGTCCACGGGTCCAAGGATGCTGGATCCAAGGATCTCTCTT  1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1561  TGGTTTCCTGCTCAGATCCAGTCCACGGGTCCAAGGATGCTGGATCCAAGGATCTCTCTT  1620

Query  1621  CCACAGTGGCCGGCCACTCAATCCCAGGGCCATTGGCCTCACCCTCCCACCCCATGTAA  1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1621  CCACAGTGGCCGGCCACTCAATCCCAGGGCCATTGGCCTCACCCTCCCACCCCATGTAA  1680

Query  1681  ATATTACTCTGTCCTCTGGGGGGCGCTCTAGGGAGCCCCTTGTGTAGATGCTCTTTAAAT  1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1681  ATATTACTCTGTCCTCTGGGGGGCGCTCTAGGGAGCCCCTTGTGTAGATGCTCTTTAAAT  1740

Query  1741  AATAAAGGTGGTTTTGATTAATGGGAGAAAAAAAAAAAAAA  1781
              ||||||||||||||||||||||||||||||||||||||||
Sbjct  1741  AATAAAGGTGGTTTTGATTAATGGGAGAAAAAAAAAAAAAA  1781
```

```
CPU time:      0.02 user secs.       0.01 sys. secs          0.03 total secs.

Lambda      K      H
    1.33    0.621     1.12

Gapped
Lambda      K      H
    1.33    0.621     1.12


  Matrix: blastn matrix:1 -2
  Gap Penalties: Existence: 5, Extension: 2
  Number of Sequences: 1
  Number of Hits to DB: 517
  Number of extensions: 20
  Number of successful extensions: 1
  Number of sequences better than 10.0: 1
  Number of HSP's gapped: 1
  Number of HSP's successfully gapped: 1
  Length of query: 1781
  Length of database: 17,904,826,729
  Length adjustment: 27
  Effective length of query: 1754
  Effective length of database: 17,904,826,702
  Effective search space: 31405066035308
  Effective search space used: 31405066035308
  X1: 11 (21.1 bits)
  X2: 26 (50.0 bits)
  X3: 26 (50.0 bits)
  S1: 14 (27.6 bits)
  S2: 22 (43.0 bits)
```

Exhibit 3. BLAST alignment of rat and mouse hepsin coding sequences



NCBI **Blast 2 Sequences results**

Protein | Entrez | BLAST | OMIM | Taxonomy | Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match 1 Mismatch -2 gap open 5 gap extension 2
x_dropoff 50 expect 10.0000 wordsize 11 Filter ☐ View options Standard
Masking character options X for protein, n for nucleotide Masking color options Black
☐ Show CDS translation Align

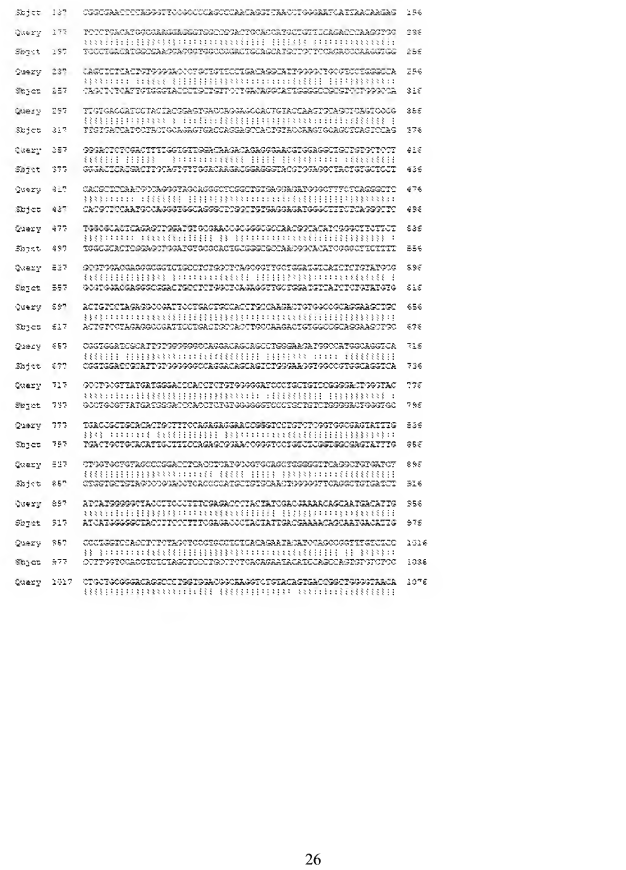
Sequence 1: gi|8395559|ref|NM_017112.1|Rattus norvegicus hepsin (Hpn), mRNA
Length = 1739 (1 .. 1739)

Sequence 2: gi|6680266|ref|NM_008281.1|Mus musculus hepsin (Hpn), mRNA
Length = 1781 (1 .. 1781)

NOTE: Bitscore and expect value are calculated based on the size of the nr database.

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

```
 Score = 2788 bits (1435),  Expect = 0.0
 Identities = 1646/1744 (94%), Gaps = 6/1744 (0%)
 Strand=Plus/Plus

Query  1     GCAGGCCCCAAGCTTGCTGGCTGCTGCTGCCACCCTTCTCCTCCCCGGGCTGCCCGCTGCTGT  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||| |||||| ||||||
Sbjct  17    GCAGGCCCCAAGCTGCTGGCTRCTRCTRCCACCCTTGGCTTCTCCRGGCTCTCCCGCTGCTGT  76

Query  61    GGGGACA----CCATGCCCCTTCCCAGGGCCGGGAGACTAACCCCAAACCTGCACCATCTC  116
              |||||     ||||||||| ||||||||||||||||||||||||||||||||||||||||
Sbjct  77    GGGGACAGAGACCATGCCCCTGCCCAGGCCCGGGAGACTAACCCCAAACCTGCACCATCTC  136

Query  117   CCCCGAACCCCAGGGTTCCGCCCCAGTCCAACAGGCCAACCTGGGAATCATTAACAAAAG  176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```



Sbjct 1037 CTTGTTGGGGGAGAGGGCCTGGTGGATGGTAAGTTGTGTATGTGACGGGCTGGGGTAATA 1096
Query 1077 CAGAGTTATATGGCAAGAAGCTGTGGTGCTCCAAAAGGCCCGGGGCCCCCATATAAGCA 1136
Sbjct 1097 CAGAGTTCTATGGCAAAAAGCTATGGTGCTCCAAAAGGCCCGTGTCTACCATAAGCA 1154
Query 1137 ACGAAGTTTGCAACAGCCCGGACTTCTACTGTAAATTCAGATCGAAGCCAAGATGTTCTGTG 1196
Sbjct 1155 AGGAAGTTTGCAACAGTCCTAGTTGTAGGGAATTGATCAAGTCTAAGATGTTCTGTG 1214
Query 1197 CTGGTCATCCAGGGTGGTATTGATGCAGCCAGGTTAACAGTGAGCCACTTTGTAC 1254
Sbjct 1215 CTTGTCATCCTGAGGTGGCATTGATGCTGCAGGCTAACAGTGAGCCACTTTGTAC 1272
Query 1255 GTGAGAACAGAATCTGTGGAACATCAAATGGGAGCTGTGGGCATTGTAAGTTGTTTA 1312
Sbjct 1273 GTGAAAACAGCATCTGTGGAAACATCAAGGTGCTGGCATAGTGTCATTATAAGCTGGGGTA 1332
Query 1313 GGGTGTGTTTGGGCGGGAAGCGGTGAAGTGTACAGCAAAGTCATTGACTTCGGGAGT 1370
Sbjct 1333 TGGTGTGTTTGGGCGGGAAGCGGTGAAGTGTACAGCAAAGTCATTGACTTCGGGAGT 1390
Query 1371 GGACTTTCTAGGTGATAAAGAGTCACTCCGAAGTCAGTGGCATGTGACTCAGCCTGAT 1430
Sbjct 1391 GGACTTTCTAGGTCATAAAAAGTCACTCCGAAGTCAGTGGCATGTGACTCAGCCTGAT 1450
Query 1431 CCCGCCTCTATCGC-CTGTTCTGTGCTGTCCAGCATCCAGAGTCAGAGTTGGTCTGGTGG 1489
Sbjct 1451 GCCGCCTCTATCTGCTGTTCTGTGCTGTCCAGCATCCAGAGTCAGAGTTGGTCTGGTGG 1510
Query 1490 CTCAGGGGTGAGTGGCAGGCTCCCACGTGGGGTATAGATGGAACGGTTTCTGCTCTGAA 1549
Sbjct 1511 CTCAGGGGTGAGTGGCAGGCTCCCACGTGGGGTCCAAGGACTTTCTTCCACAATGGCGGCCC 1570
Query 1550 TCCATCTATAATCCAAGAATGCTGGGTCCAAGGACTTTCTTCCACAATGGCGGCCC 1603
Sbjct 1571 TCCATCTACGGGTCCAAGGATGCTGGATCCAAGTATTTGTGTCCACAGTGGCGGCCC 1626
Query 1604 AGTCAATGCCAGGGCCATTGTGCTGACCTCCACGCCATGTAAATATTACTCTGTCCTC 1663
Sbjct 1627 AGTCAATGCCAGGGCCATTGTGCTGACCTCCACGCCATGTAAATATTACTCTGTCCTC 1686
Query 1664 TGGGGGTTGTTTTGGAGGGAGGGTTTGTGCGGATGCTCTTAAATAATAAAGGTGGTTTT 1723
Sbjct 1687 TGGGGGGTGCTGTAG-GGAGCCCCTTGTGCAGATGCTCTTAAATAATAAAGGTGGTTTT 1745
Query 1724 GATT 1728
Sbjct 1746 GATT 1750

```
CPU time:      0.02 user secs.       0.00 sys. secs          0.02 total secs.

Lambda     K      H
    1.33    0.621     1.12

Gapped
Lambda     K      H
    1.33    0.621     1.12


  Matrix: blastn matrix:1 -2
  Gap Penalties: Existence: 5, Extension: 2
  Number of Sequences: 1
  Number of Hits to DB: 331
  Number of extensions: 7
  Number of successful extensions: 4
  Number of sequences better than 10.0: 1
  Number of HSP's gapped: 1
  Number of HSP's successfully gapped: 1
  Length of query: 1739
  Length of database: 18,244,403,944
  Length adjustment: 27
  Effective length of query: 1712
  Effective length of database: 18,244,403,917
  Effective search space: 31234419808904
  Effective search space used: 91734419808904
  T1: 21 (21.1 bits)
  T2: 26 (50.0 bits)
  X1: 26 (50.0 bits)
  S1: 14 (27.4 bits)
  S2: 22 (43.0 bits)
```

Exhibit 4. BLAST alignment of SEQ ID NO:1 with BC025716.1

NCBI **Blast 2 Sequences results**

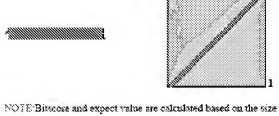
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 10.0000 wordsize: 11 Filter: ☐ View option: Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
☐ Show CDS translation Align

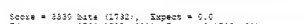
Sequence 1. lcl|seq_1
Length = 1783 (1 .. 1783)

Sequence 2: gi|19343934|gb|BC025716.1|Homo sapiens hepsin (transmembrane protease, serine 1), mRNA (cDNA clone MGC:34588 IMAGE:5228523), complete cds
Length = 1761 (1 .. 1761)

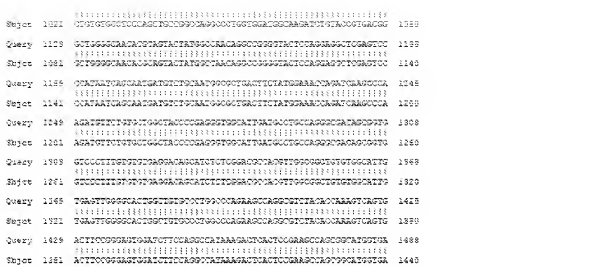


NOTE:Bitscore and expect value are calculated based on the size of the nr database.

NOTE If protein translation is reversed, please repeat the search with reverse strand of the query sequence.



```
 Score = 3330 bits (1732),  Expect = 0.0
 Identities = 1794/1798 (99%), Gaps = 0/1798 (0%)
 Strand=Plus/Plus

Query  49    GCCTGGCCTAGCAGGGCCCAGGCCACCGCCTCTGCCTCCAGGCGCCCCGCTGCTGCGGGG  108
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     GCCTGGCCTAGCAGGGCCCAGGCCACCGCCTCTGCCTCCAGGCGCCCCGCTGCTGCGGGG  60

Query  109   CCACCATGCCCCTGCCCAGGCCTGAAGACTGACCCGACCCCGCCATCTCCTGAGGCCTCC  168
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    CCACCATGCCCCTGCCCAGGCCTGGAGACTGACCCGACCCCGCCATCTCCTGAGGCCTCC  120

Query  169   GCCCCCACCCTCTGGACCCCAGGGTCCCACCCTGCCCAGGAGTCACCCAGGAATCAT  224
```



```
CPU time:     0.62 user secs.      0.00 sys. secs          0.62 total secs.

Lambda      K      H
    1.33   0.621     1.12

Gapped
Lambda      K      H
    1.33   0.621     1.12

Matrix: blastn matrix:1 -2


Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 1
Number of Hits to DB: 488
Number of extensions: 8
Number of successful extensions: 1
Number of sequences better than 10.0: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Length of query: 1781
Length of database: 15,244,403,944
Length adjustment: 27
Effective length of query: 1754
Effective length of database: 15,244,403,917
Effective search space: 26737178273542
Effective search space used: 26737178278262
X1: 11 (21.1 bits)
X2: 26 (50.0 bits)
X3: 26 (50.0 bits)
S1: 14 (27.6 bits)
S2: 27 (49.0 bits)
```

Exhibit 5. BLAST alignment of SEQ ID NO:1 with NM_182983.1

NCBI **Blast 2 Sequences results**

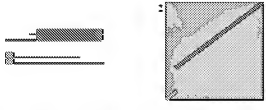
Nucleotide Protein BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match 1 Mismatch -2 gap open 5 gap extension 2
x_dropoff 50 expect 0.05 wordsize 11 Filter View option Standard
Masking character option X for protein, n for nucleotide Masking color option Black
Show CDS translation Align

Sequence 1 lcl|seq_1
Length = 1783 (1 .. 1783)

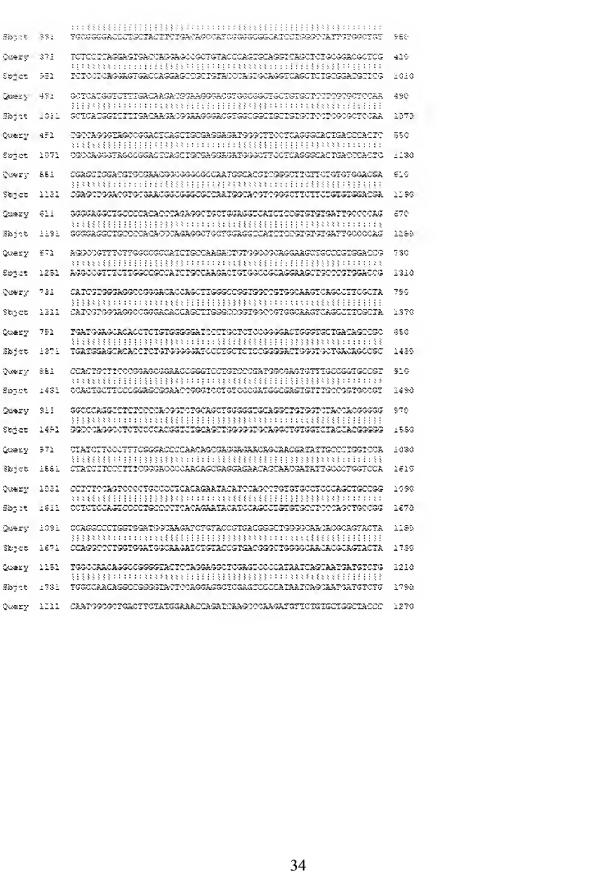
Sequence 2 gi|32699513|ref|NM_182983.1|Homo sapiens hepsin (transmembrane protease, serine 1) (HPN), transcript variant 1, mRNA
Length = 2363 (1 .. 2363)

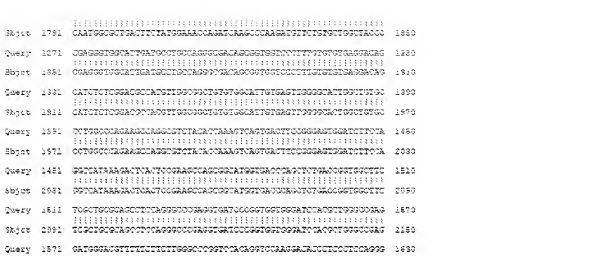


NOTE Bitscore and expect value are calculated based on the size of the nr database.

NOTE If protein translation is reversed, please repeat the search with reverse strand of the query sequence

```
 Score = 2063 bits (1059),  Expect = 0.0
 Identities = 1059/1059 (100%), Gaps = 0/1059 (0%)
 Strand=Plus/Plus

Query  191   GTTCCCACCCTGGCCAGGAGTTCAGCAGGGATCATTACAAGAGTCATGACATGGC  250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  771   GTTCCCACCCTGGCCAGGAGTTCAGCAGGGATCATTACAAGAGTCATGACATGGC  830

Query  251   GCAAAAGGAAGGTGGCCGGACTGTCCCATGCTGCTGCAGACCCAAGGTGGCAGTCTCCAC  310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  831   GCAAAAGGAAGGTGGCCGGACTGTCCCATGCTGCTCCAGACCCAAGGTGGCAGTCTCCAC  890

Query  311   TGCGGGGACCCTGCTACTCTGACAGCCATCCGGGCCGGCATCCTGGGCAATTGTGGCTGT  370
```



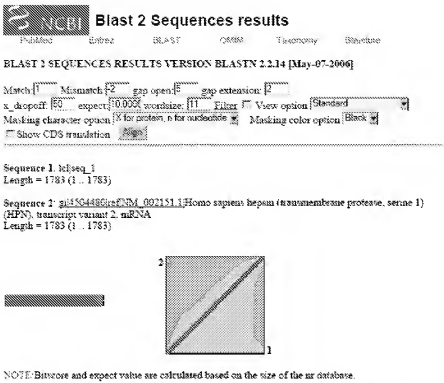


```
Sbjct  1791  CAATGGCGCTJACTTCTADGGAAADAAATJAAGCCCAAGATJTTCTGCTGTCTAOCC  1850

Query  1271  CGGGGGTGGTATTGA2OOCTGCCAGGGGACAGGGGTGCTCCCTTTCTGTGTCAGGAG  1330

Sbjct  1851  CAAGGGTGGCACTGATGCCTGCCAGGTCACAGGGGTGCTGGTCCCTTTGTGTGTCAGGAG  1910

Query  1331  CATCTCTCGGACGCCADTTDGCGGCTGTGTGGCATTGTGAGTTGGGCATTGTGTGTGGC  1390

Sbjct  1911  CATCTCTCGGACGCCADTTDGCGGCTGTGTGGCATTGTGAGTTGGGCATTGTGTGTGGC  1970

Query  1391  CTTGGCCAGAAAGCCAGGCGTCTACACTAAAGTCAGTGACTTCGGGGAGTGGATCTTCTA  1450

Sbjct  1971  CTTGGCCAGAAGCCAGGCGTCTACACTAAAGTCAGTGACTTCGGGGAGTGGATCTTCTA  2030

Query  1451  GGCCATAAAGACTCATCCGAAGCCAGCGGCATGGTGACCCACTTGGAGCGGTGGCTTC  1510

Sbjct  2031  GGCCATAAAGAATCACTCGGAAGCCAGCGGCATGGTGACCCACTTGGAGCGGTGGCTTC  2090

Query  1511  TCGCTGGCAGCCTCCAGGGCCCAAGTGATCGGGGTGGTGGGATCTACTCTGGGCCAAG  1570

Sbjct  2091  TCGCTGGCAGCCTCCAGGGCCCAAGCAAGCGGGTGGTGGGATCTACCCTGGGCCAAG  2150

Query  1571  GATGGAACGTTTTCCTTCTTGGGCCCGGTCCAGAGGTCCAAGAACACTTCCTCCTTCAGGG  1630

Sbjct  2151  GATGGGACGTTTTCCTTCTTGGGCCCGGTCCAGAGGTCCAAGGACACTTCCTCCTTCAGGG  2210

Query  1631  TCCTTCTTCTTGAGAGTGGCGGGCCCACCTAGCCCCGAGAGTCAJCAAAGCTCAGCCTCCTG  1690

Sbjct  2211  TCCTTCTTCTTGAGAGTGGCGGGCCCACCTAGTCCCGAGAGTCACCAAACTCAGCCTCCTG  2270

Query  1691  AOCCCCATGTAAATATTGTTTTGCTGTCTGGGAGTCCTGTCTAGGTGCCCCTGATGATGG  1750

Sbjct  2271  ACCCCCATGTAAATATTGTTTTGCTGTCTGGGAGTCCTTGTCTAGGTGCCCCTGATGATGG  2330

Query  1751  TATGCTTTTAAATAATAAAGATGGTTTTGATT  1783

Sbjct  2331  GATGCTCTTAAATAATAAAGATGTTTTTGATT  2363
```

```
Score =  371 bits (193),  Expect = 4e-99
Identities = 193/193 (100%), Gaps = 0/193 (0%)
Strand=Plus/Plus

Query  1   TGGAGCCGGCTTTCCAGGGACCCTAGCTAGGGCCCACAGGTGAGGCAGCCTGGCCTAGC  60

Sbjct  1   TCGAGCCGGCTTTCCAGGGACCCTAGCTAGGGCCCACAGGTGAGGCAGCCTGGCCTAGC  60

Query  61  AGGCCCCAGGCCACGGCTCTGCCTCCAGGCCGCCCGCTGTGCGGGGCCACCATGCTCC  120

Sbjct  61  AGGCCCCAGGCCACCGCCTCTGCCTCCAGGCCGCCCGCTGTGCGGGGCCACCATGCTCC  120

Query  121 TGCCCACGCCTGAAGACTGACCCGACCCCGCCATACCTCTGAAGAATCCCCCCCCAGCTGC  180

Sbjct  121 TGCCCAGGCCTGAAGACTGACCCGACCCCGCCATACCTCTGAAGAATCCCCCCCAGCTGC  180
```

```
Query  191  TGGAGCCCAAGGTT  193
            ||||||||||||||
Sbjct  191  TGGAGCCCAAGGTT  193


CPU time:     0.02 user secs.      0.00 sys. secs          0.02 total secs.

Lambda     K      H
    1.33   0.621     1.12

Gapped
Lambda     K      H
    1.33   0.621     1.12


Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 1
Number of Hits to DB: 820
Number of extensions: 7
Number of successful extensions: 2
Number of sequences better than 10.0: 1
Number of HSP's gapped: 2
Number of HSP's successfully gapped: 2
Length of query: 1758
Length of database: 18,246,408,904
Length adjustment: 57
Effective length of query: 1758
Effective length of database: 18,246,408,847
Effective search space: 32093172278252
Effective search space used: 32093172278252
X1: 11  (21.2 bits)
X2: 26  (50.0 bits)
X3: 26  (50.0 bits)
S1: 14  (27.6 bits)
S2: 22  (43.0 bits)
```

Exhibit 6. BLAST alignment of SEQ ID NO:1 with NM_002151.1



NCBI **Blast 2 Sequences results**

Publlled Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match 1 Mismatch -2 gap open 5 gap extension 2

x_dropoff 50 expect 10.0000 wordsize 11 Filter Vsew option Standard

Masking character option X for protein, n for nucleotide Masking color option Black

Show CDS translation Align

Sequence 1: lcl|seq_1
Length = 1783 (1 .. 1783)

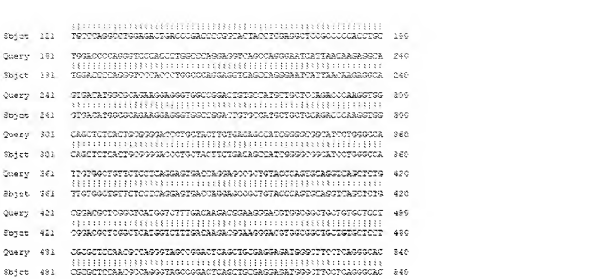
Sequence 2: gi|45644866|ref|NM_002151.1|Homo sapiens hepsin (transmembrane protease, serine 1) (HPN), transcript variant 2, mRNA
Length = 1783 (1 .. 1783)

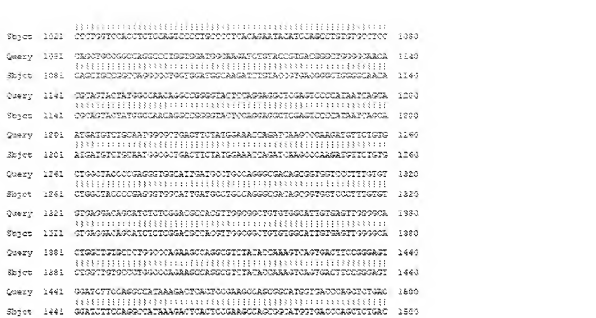
NOTE: Bitscore and expect value are calculated based on the size of the nr database.

NOTE: If protein translation is reversed, please repeat the search with a reverse strand of the query sequence

Score = 3429 bits (1783), Expect = 0.0
Identities = 1783/1783 (100%), Gaps = 0/1783 (0%)
Strand=Plus/Plus

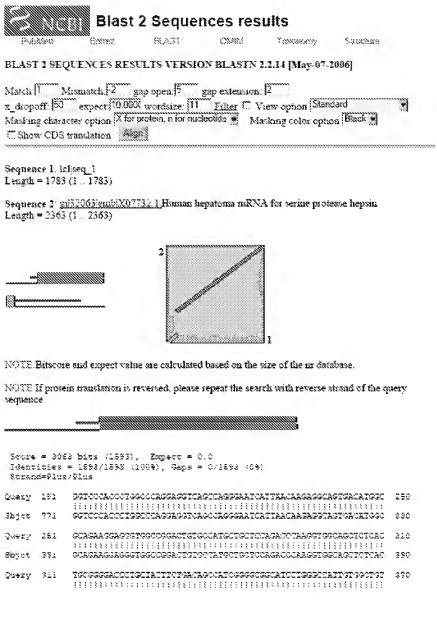
```
Query  1     TCGAGCCCGCGTTTCCAGGGACCCTACCTGAGGGCCCACAGGTGAGGCAGCCTGGCCTAGC  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     TCGAGCCCGCGTTTCCAGGGACCCTACCTGAGGGCCCACAGGTGAGGCAGCCTGGCCTAGC  60

Query  61    AGGCCCCACGCCACCGCCTCTGCCTCAGGCGGCCCCTGCTGCTGCGGGACCACCATGCTCC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    AGGCCCCACGCCACCGCCTCTGCCTCAGGCGGCCCCTGCTGCTGCGGGACCACCATGCTCC  120

Query  121   TGCCCAGGCCTGGAGACTGACCGAGCCCCGCACTACCTCGAGGCTCCCCCCCCACCTGC  180
```





```
CPU time:      0.02 user secs.       0.01 sys. secs          0.03 total secs.

Lambda      K      H
    1.33   0.621     1.12

Gapped
Lambda      K      H

    1.33   0.621     1.12


Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 1
Number of Hits to DB: 808
Number of extensions: 5
Number of successful extensions: 1
Number of sequences better than 10.0: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Length of query: 1795
Length of database: 13,244,408,944
Length adjustment: 27
Effective length of query: 1766
Effective length of database: 13,244,408,917
Effective search space: 23387178278252
Effective search space used: 23387178278252
X1: 11 (21.1 bits)
X2: 38 (50.0 bits)
X3: 96 (95.0 bits)
S1: 19 (27.6 bits)
S2: 22 (43.0 bits)
```

Exhibit 7. BLAST alignment of SEQ ID NO:1 with X07732.1



NCBI **Blast 2 Sequences results**

Nucleotide Protein BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match 1 Mismatch -2 gap open 5 gap extension 2
x_dropoff 50 expect 10.000 wordsize 11 Filter View option Standard
Masking character option X for protein, n for nucleotide Masking color option Black
Show CDS translation Align

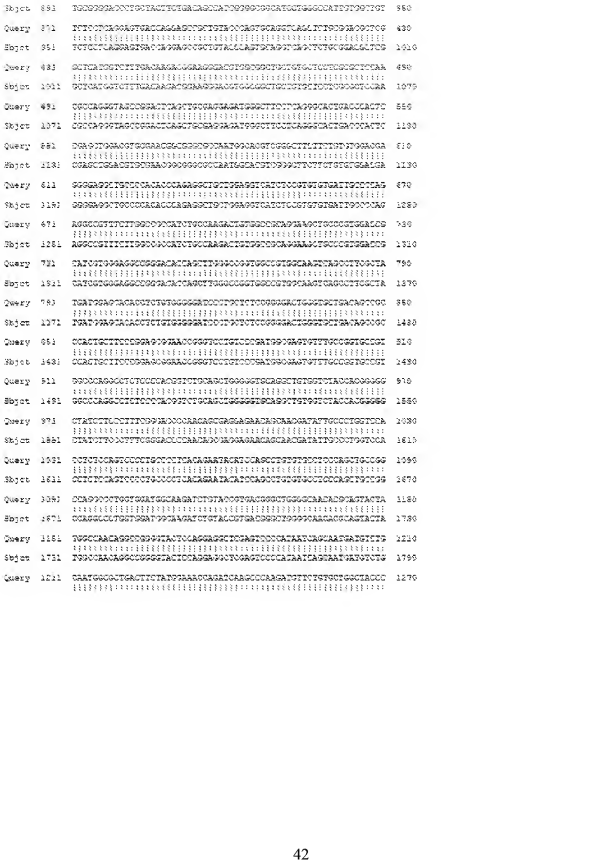
Sequence 1: lcl|seq_1
Length = 1783 (1 .. 1783)

Sequence 2: gi|72563|emb|X07732.1|Human hepatoma mRNA for serine protease hepsin
Length = 2363 (1 .. 2363)

NOTE Bitscore and expect value are calculated based on the size of the nr database.

NOTE If protein translation is reversed, please repeat the search with reverse strand of the query sequence

```
 Score = 3069 bits (1593),  Expect = 0.0
 Identities = 1593/1593 (100%), Gaps = 0/1593 (0%)
 Strand=Plus/Plus

Query  191   GGTCCCAGCCTGGCCAGGAGTCAGCCAGGAATCATTACAAGAGGCAGTGACATGGC  250
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  771   GGTCCCAGCCTGGCCAGGAGTCAGCCAGGAATCATTACAAGAGGCAGTGACATGGC  830

Query  251   GCAGAAGGAGGTGTGGCCGGATGTGCCATGCCTGCTCAGACTCAAGTCGGCAGTCTCAC  310
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  831   GCAGAAGGAGGTGGCCGGATGTGCCATGCCTGCTCAGACTCAAGTCGGCAGTCTCAC  890

Query  311   TGCGGGACCTTGCTACTTCTGACAGCACCGGGCGGCATCTTGGGCATTTGTGGCTGT  370
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```



Sbjct 891 TGCGGGGAGCCTTGCTACCTGTGACAGCCACTCGGGGCGGCATGGCTGGGGCATTCTTGCTGT 950

Query 871 ATTCCTCAGGAGTACCAGAGCTGCTGTACCAGTGCAGCTCAGLATTCGGGGAGCCTCG 930
::
Sbjct 951 TCTCCTCAGGAGTACCAGGAGCCGCTGTACLAGTGCAGCTCAGATCAGCTTTGCGGACLCTCG 1010

Query 931 GCTCATCGCTCTTTGACAAGALGGAAGGGACCTGGCGGCTGTGTGCLTCCTCGGCTCCAA 990
::
Sbjct 1011 GCTCATGGTCTTTGACAACACGGAAGGAAGTGGLGGCTGCTGCTGCTCCTCGGCTCCAA 1070

Query 991 CGCCAGGGTAACCGGATTACTTGCGAGAAATGGGCTTCTTAGGGACTGACCLCACTC 1050
::
Sbjct 1071 CGCCATGGTACTGATCGGATTACTTGCGAGAAATGGGCTTCTTAGGGACTGACCTCACTC 1130

Query 1051 CGATGTGAAGGTGGAAACGGCGGGCGTCAATGGAGTGGGCGTTLTTTTGTGGAAGA 1110
::
Sbjct 1131 CGAGCCGAACGTCGTAAGTGGGGGGTCAATGGCACTCTGGGTTTCTTGTTGCGAACGA 1190

Query 1111 GGGAAGTCTGTCCCACACCCAAAGGCTGTTGGAGTCATCTCCGTGTGTAATTGTCCTAG 1170
::
Sbjct 1191 GGGAAGTCTGCCCCACACCCAGAGGCTGTTGGAGTCATCTCCGTGTGTAATTGTCCTAG 1250

Query 1171 AGGCCGTTCTTTGCCCCCATCTGCCAAGACTATGCCTCAGTAGTAAAGTCCCGTGGACCG 1230
::
Sbjct 1251 AGGCCGTTCTTTGCCCCCATCTGCCAAGACTATGCCTCAGTAGTAAAGTCCCGTGGACCG 1310

Query 1231 CATCGTGGAGGCCGGGACACCAGCTTGGCCGGTGGCCGTGGCAATCAGCGTTGGTA 1290
::
Sbjct 1311 CATCGTGGAGGCCGGGACACCAGCTTGGCCGGTGGCCGTGGCAATCAGCGTTGGCTA 1370

Query 1291 TGATGGAGTACACGCTGTTGTGGGGGATCCTTCTCCTGGGGAGCTGGTGCTGACAGCCGC 1350
::
Sbjct 1371 TGATGGAGTACACCTCTGTGGGGGATCCTTCTCCTGGGGAGCCGGTGCTGACAGCCGC 1430

Query 1351 CCACTGCTTCCCGGAGYGAAGCGGGTCCTGCTCCGATGGYGAGTGTTGCCGGTGCCGT 1410
::
Sbjct 1431 CCACTGCTTCCCGGAGCGAAGCGGGTCCTGCTCCGATGGCGAGTGTTGCCGGTGCCGT 1490

Query 1411 GGCCCAGGCCTGTCCCACGTCTCGAGTGGGGGTCGCGCCTGTGGTCIACGCGGGGG 1470
::
Sbjct 1491 GGCCCAGGCCTCTCCCACGTCTCGAGTCGGGGGTCACGCCTGTGGTCTACGCAGGGGG 1550

Query 1471 CTATCCTCCCTTTCGGGACCCAAGAGCGAGGAGAACGCAATGCAATGTCGCCCTGGTCCA 1530
::
Sbjct 1551 CTATCTTCCCTTTCGGGACCCAAGAGCGAGGAGAACGCAACGATATTGCCCTGGTCCA 1610

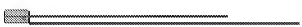
Query 1531 CCTCGTCGAGTCCCCTGCCCCTCACAGAATACGTCCAGCCTGTTTCCCCAGCTGLCGG 1590
::
Sbjct 1611 CCTCTCGATCCCCTGCCCCTCACAGAATACATCCAGCCTGTGTTGCCTCCAGCTGTCCG 1670

Query 1591 CCAGGCCTGGTGAATGGCAAGATGTTATACGTGACGGGCTGGGGCGACAACGGCAGTACTA 1650
::
Sbjct 1671 CCAGGCCCTGGTGAATGGCAAGATGTTGACGTGACGGGCTGGGGCGACAACGGCAGTACTA 1730

Query 1651 TGGCCAACAGGCCGGGTTACTCCAGGAGGCTCGATCCCCACATAATCAGCAATGACTTTT 1710
::
Sbjct 1731 TGGCCAACAGGCCGGGTTACTCCAGGAGGCTCGACCCCCACAATCAGCAATGACTTTTT 1790

Query 1711 CAATGGGGATGACTTCTATGTAAACAGATCAAGCCCAAGACTTTCTGTGCTGGCTAGCCC 1770
::
Sbjct 1791 TTGCCAAAAGCCCGGGGTACTCCAGGAGGCTCCCCACAATCAGCAATGACTGTTTTGCCC 1850

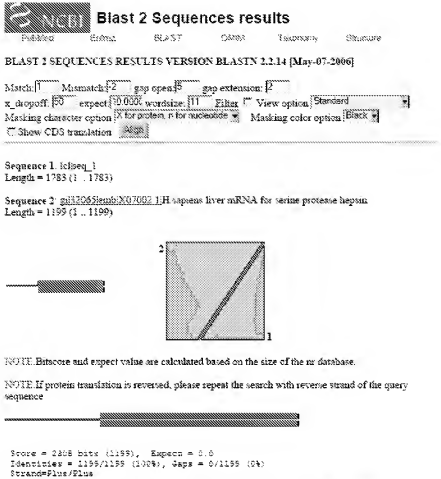
```
Sbjct  1781  CAATGGGGCTGACTTGTATGTGAAAGCAAATCAAGCCCAAGATGTGTGTATGGCTGGCCAGCC  1850

Query  1271  CAAGGGTGGCATTATGGTTGGAAGGGGACAGCTGGTGGTCCTTTGTGTGTGAGGACAG  1330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1841  CAAGTGTGGCACTAATGGCCTGCCAGTGGCAGAGGTGACCCCTTCTTGTGTGAGGACAG  1910

Query  1331  CATCTCTGGGACGCCACGTTGTGGGTGTGGCATTGTGAGTTGTGGGACTGTGGCTGCC  1390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1911  CATCTTTGGGATGCCACGTTGGCGGCTGTGTGGTATTGTGAGTTGGGGCACTGGCTGTGC  1970

Query  1391  CCTGGTGGCAAAGCCAGGGCGTCTACACTAAATGTCATGACTTCCGGGAGTGTATCTTCCA  1450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1971  CCTGGGCCAGAAGCCAGGGTGTACAGCAAAGTCAGTGACTTCCGGGAGTGGATCTTCCA  2030

Query  1451  GGCCATAAAGACTCACTCCAAAGCCAGCGGCATGGTAGCCAGCTGTGACCGGTGGTTC  1510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2031  GGCCATAAAGACTCACTCCAAAGCCAGCGGCATGGTGACCCAGCTCTGACCGGTGGTTC  2090

Query  1511  TGGATGGGAGCCTCCAGGGGGTGAGGTGATCCGGTGGTGGATCCAGCGTGGGCCAG  1570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2091  TGGTGGGAGCCTCCAGGGGGCAGGTGATCCCGGTGGTGGATCCAGCGTGGGCCAG  2150

Query  1571  GATGGAAGTTTTTCTTGTGGGCCTGGTCAGAGGTCCAAGGACACTTCCTTCCAGGG  1630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2151  GATGGAGGTTTTTCTTGTGGGCCTGGTCAGAGGTCCAAGGACACTTCCTTCCAGGG  2210

Query  1631  TCGTCTTTTCAGAGTGGGGGCCACTCAGCCCCAGACCAGCCAAGCTCAGCCTCCTG  1690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2211  TCGTCTTTTCAGAGTGGGGGCCACTTAGCCCCAGACCAGCCAAGCTCAGCCTCCTG  2270

Query  1691  ACCCCGATGTAAATATTGTTCTATCTGTGGGACTCCTTATTAGTCCCCTGATGACGG  1750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2271  ACCCCGATGTAAATATTGTTCTGCTGTCTGGGACTCCTGTCTATGTGCCCCTGATGATGG  2330

Query  1751  GATGGTCTTTAAATAATAAAGATGTTTTTGATT  1782
              ||||||||||||||||||||||||||||||||
Sbjct  2331  GATGGTCTTTAAATAATAAAGATGTTTTTGATT  2362
```



```
Score =  371 bits (193),  Expect = 4e-98
Identities = 193/193 (100%), Gaps = 0/193 (0%)
Strand=Plus/Plus

Query  1    TCAAGCCCAGGTTTCCCAGGGACCCTACCTGAGGGCCACAGGTGACCACAGGTGACCAGCTCAGCC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    TCAAGCCCAGGTTTCCCAGGGACCCTACCTGAGGGCCCACAGGTGACCACAGGTGACCAGCTCAGCC  60

Query  61   ACGGCCCAAGGCCACGGCCTCTGCCTCCAGGGCTGCCCACTGCTGCCGGGCTACCAIGCTCC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   ACGGCCCAAGGCCACGGCCTCTGCCTCCAGGGCTGCCCACTGCTGCCGGGCTACCAIGCTCC  120

Query  121  TGCCTAGGCTTGAAAGCTGACCGGACCCCGGCATTATGTTGAGGCTCCCCACCTTC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  TGCCTAGGCTTGAAAGCTGACCGGACCCCGGCATTATGTTGAGGCTCCCCACCTTC  180
```

```
Query  181  TGGAGCCCAAGGT  193
            ||||||||||||
Sbjct  181  TGGAGCCCAAGGT  193


CPU time:      0.02 user secs.       0.00 sys. secs          0.02 total secs.

Lambda     K      H
    1.33    0.621     1.12

Gapped
Lambda     K      H
    1.33    0.621     1.12


Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 1
Number of Hits to DB: 820
Number of extensions: 7
Number of successful extensions: 2
Number of sequences better than 10.0: 1
Number of HSP's gapped: 2
Number of HSP's successfully gapped: 2
Length of query: 1788
Length of database: 18,244,403,924
Length adjustment: 27
Effective length of query: 1786
Effective length of database: 18,244,403,817
Effective search space: 32587179278262
Effective search space used: 32587179278262
X1: 12 (21.1 bits)
X2: 25 (50.0 bits)
X3: 25 (50.0 bits)
S1: 14 (27.6 bits)
S2: 22 (43.0 bits)
```

Exhibit 8. BLAST alignment of SEQ ID NO:1 with X07002.1



NCBI **Blast 2 Sequences results**

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 10.0000 wordsize: 11 Filter: View option: Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
Show CDS translation Align

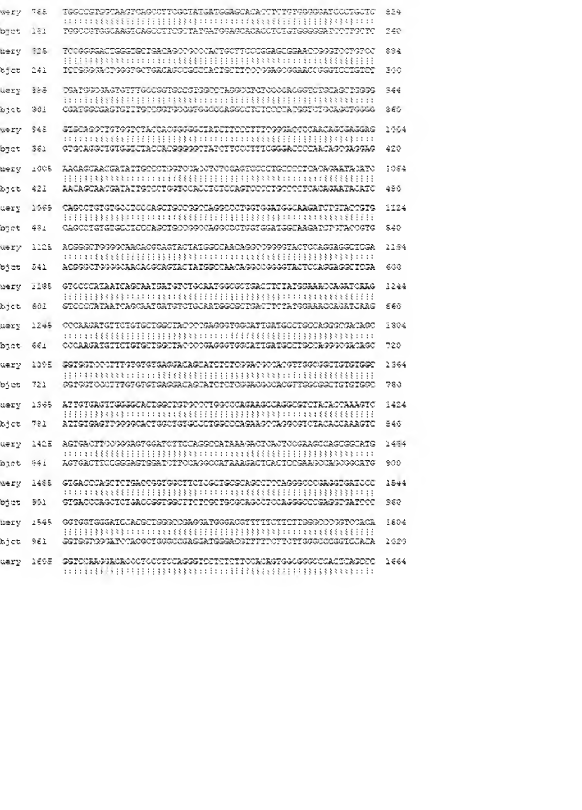
Sequence 1: lcl|seq_1
Length = 1783 (1 .. 1783)

Sequence 2: gi|4206|emb|X07002.1|H.sapiens liver mRNA for serine protease heptsin
Length = 1199 (1 .. 1199)

NOTE: Bitscore and expect value are calculated based on the size of the nr database.

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence

```
 Score = 2208 bits (1195),  Expect = 0.0
 Identities = 1199/1199 (100%), Gaps = 0/1199 (0%)
 Strand=Plus/Plus

Query  585   ACGTCGGGGTTCTTCTGTGTGCGGACGAGCGGAAGATGGCCCACACCCAGAGGCTGCTGAG  644
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1     ACGTCGGGGTTCTTCTGTGTGCGGACGAGCGGAAGATGGCCCACACCCAGAGGCTGCTGAG  60

Query  645   GTCATCTCCTTGTGTGTATTGCCCAGGGCCGTTTCTTGGCCCCATCTGCCAAGACTGT  704
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61    GTCATCTCGTTGTGTATTGCCCAGAGGCCGTTTCTTGGCCCCATCTGCCAAGACTGT  120

Query  705   GGCTCCAGTAAGTTCCCCTCTAAGCGATGCCCGGAGGTGGTTCAGCAGCTGCCCCG  764
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```



Sbjct 121 GGCGCGAGGAAGTIGCCCTFTGAACGAGATCGTGGGAGGCCGGGAGTAAAAGTTGGGCGG 180

Query 765 TGGCGGTGTGCAAGTGAGCTTTGGGTATGATGGAGCACACTTCTTTGTGGAATCCGTAGCTG 824

Sbjct 181 TGGCGGTGGGAAGTGAGCTTCGTATATGTGAACAACACTGTGTGGGGGATCCCTTCCTC 240

Query 825 TCCGGGGAGCTGGGGTGTGACAGCCACCTATGGTTCCCGGAGCGGAACCGGGTCCTGTCC 884

Sbjct 241 ICCGGGAGCTGGGGTGTGACAGCGGCCACTGCTTCCCTGGGCGGAACCGGGTCCTGTCC 300

Query 885 TCATGTGCTATGTTTGGCGGTGCCCTGGCTCAGGCTCTCCGCAGGGGTCTGCAGCTCTGGG 944

Sbjct 301 CCATGTGCTATGTTTGGCGGTGCCCTGGCTCAGGCTCTCCGCAGGGGTCTGCAGCTCTGGG 360

Query 945 GTGCAGCTTGTTGTCTACTACGTGGTGCTATCTTTCTTTTCGTTCCTAACAGCGAAGAG 1004

Sbjct 361 GTGCATGCTGTGGCTACTACGTGGTGCTATCTTCCTTTTCGGGAGCTCCAACAGCGAAGAG 420

Query 1005 AAGAGCAACGATATTGCCCTGTTCTACATCTCGAGTCCCTGTCCCCCTCACACAAATACATC 1064

Sbjct 421 AACAGCAACGATATTGCCCTGTTCTACATCTCGAGTCCCTGTCCCCCTCACACAAATACATC 480

Query 1065 CAGCCTGTGTGGCCCCCAGCTGCCTGCTGGGGCCTTTTGTTGAAGGCAAGATTTTTACCGTG 1124

Sbjct 481 CAGCCTGTGTGGCCCCCAGCTGCCTGCTGGGGCCTTTTGTTGAAGGCAAGATTTTTACCGTG 540

Query 1125 ACGGGGTTTGGGGCAAGAGGTGAGTACTATGGCCAACAGGCCGTGGGTACTCCTAGGAGGGTCGAA 1184

Sbjct 541 ACGGGGTTTGGGGCAACAGGCAGTACTATGGCCAACAGGCCGTGGGTACTCCTAGGAGGGTCTCA 600

Query 1185 GTCCCCATAATCAGCAATGATGTTTGCAATGGGGGTGACCTTCTATGGAAACTCAAATCAAG 1244

Sbjct 601 GTCCCCATAAATCAGCAATGATGTTTGCAATGGGGGTGACCTTCTATGGAAACTCAAATCAAG 660

Query 1245 CCCAAATGTTGTGTGCTGGCTACCCCAAGGGTGGCATTGATGCCTGCCAGGCTACAGC 1304

Sbjct 661 CCCAAGATGTTCTGTGCTGGCTACCCCAAGGGTGGCATTGATGCCTGCCAGGCTACAGC 720

Query 1305 GGTGTTCCTTTTGTGTGTGAGGACAGCATCTTCTGTAACTCCACGTTGGCGTGTGTGGC 1364

Sbjct 721 GGTGTTCCGTTTGTGTGTGAGGACAGCATCTTCTGTAACGCCACGTTGGCGACTGTGTGGC 780

Query 1365 ATTGTGATTGGGGCACTGGCTGTCCCTGGCCCAGAAGCCAGGCTCTACACCAAAGTC 1424

Sbjct 781 ATTGTGAGTCCGGGCACTGGCTGTCCCTGGCCCAGAAGCCAGGCTCTACACCAAAGTC 840

Query 1425 AGTGATCTTCCGGGAGTGGATGTTCCAGGCCATAAACTCCAATCTGCGAAGCCAGCGGCATG 1484

Sbjct 841 AGTGACTTCCGGGAGTGGATGTTCCAGGCCATAAAGACTGACTCCGCAAGCCAGCGGCATG 900

Query 1485 GTGACCCAGCTCTTACCGGTGGCTCTGGCTGCGCAGCTTCCAGGGCCCAAGGTGATCCC 1544

Sbjct 901 GTGACCCAGCTCTTACCGGTGGCTCTGGCTGCGCAGCTTCCAAGGGCCCAAGGTGATCCC 960

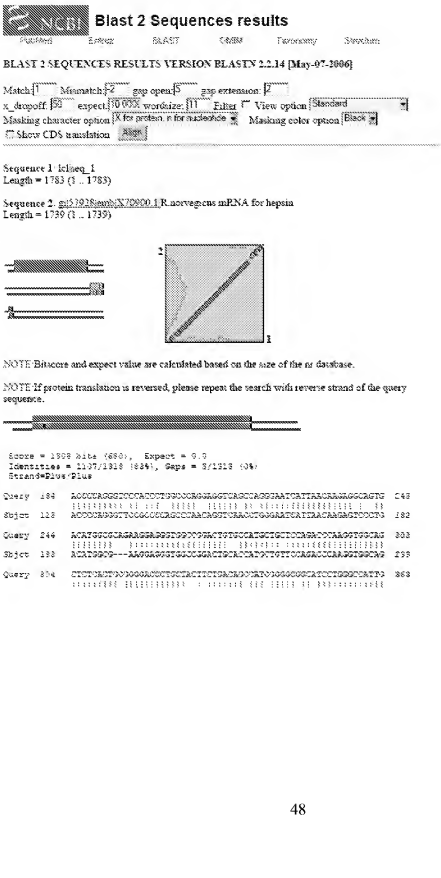
Query 1545 GGTGGTGGGATCCAGCTGGGCCGAGGATGGGAGGTTTTTCTTCTTGGGCCCGTCCACA 1604

Sbjct 961 GGTGGTGGPATCCAGCTGGGCCGAGGATGGGAGGTTTTTCTTCTTGGGCCCGTCCACA 1020

Query 1605 GGTCCAAGAGACAGCTCCTTCAGGGTCCTTCTTCCACAGTGGCGGGCCCACTCAGCCC 1664

```
Sbjct  1821  GGTGCAAGGACACCCTCCTTCTCAAGGGTCCTCCTTGCACAGTCTTTGGGGGGGAGTCAGCCC  1880

Query  1665  CAAGACCACCCAACCTCAGCCCCCTATGACCCCCATTCAAATATTCTCTGTGTGTCTGGGAC  1724
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1761  CAAAGCCACCCAACCTCACCCCCTCCACCCCCCATGCAAATATTCTCCTCTGTCTCTGGGAC  1820

Query  1725  TCCTGTCTAGGTGCCCCTCATCATGGGATGCTCTTTAAATAATAAAGATGGTTTTGATT  1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1821  TCCTGTCTAGGTGCCCCTCATCATGGGATGCTCTTTAAATAATAAAGATGGTTTTGATT  1878


CPU time:     0.02 user secs.      0.01 sys. secs          0.03 total secs.

Lambda      K      H
   1.33   0.622    1.12

Gapped
Lambda      K      H
   1.33   0.621    1.12


Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 1
Number of Hits to DB: 326
Number of extensions: 9
Number of successful extensions: 1
Number of sequences better than 10.0: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Length of query: 1753
Length of database: 18,244,403,844
Length adjustment: 27
Effective length of query: 1756
Effective length of database: 18,244,403,817
Effective search space: 32037173278252
Effective search space used: 32037173278252
X1: 11 (21.0 bits)
X2: 26 (50.0 bits)
X3: 29 (55.8 bits)
S1: 18 (35.7 bits)
S2: 22 (43.0 bits)
```

Exhibit 9. BLAST alignment of SEQ ID NO:1 with rat and mouse hepsin coding sequences



NCBI Blast 2 Sequences results

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.2.14 [May-07-2006]

Match: 1 Mismatch: -2 gap open: 5 gap extension: 2
x_dropoff: 50 expect: 10.0005 wordsize: 11 Filter: View option: Standard
Masking character option: X for protein, n for nucleotide Masking color option: Black
Show CDS translation Align

Sequence 1: lcl|seq_1
Length = 1783 (1 .. 1783)

Sequence 2: gi|57928|emb|X70900.1|R.norvegicus mRNA for hepsin
Length = 1739 (1 .. 1739)

NOTE: Bitscore and expect value are calculated based on the size of the nr database.

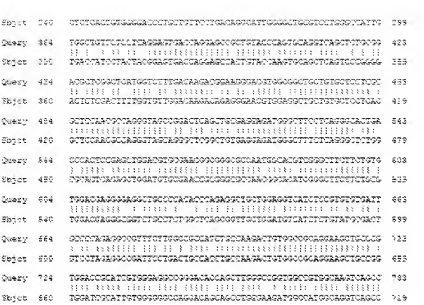
NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence.

Score = 1908 bits (960), Expect = 0.0
Identities = 1197/1918 (84%), Gaps = 9/1918 (0%)
Strand=Plus/Plus

Query 184 AGCCTAGGGCCCAGCCTGGUGCAGAAGGTCAGCCAGGGAATCATTAACAAAAGGCAGTG 243
Sbjct 128 AGCCCAGGGTCCAGCCCCAACAAGGTCAAACATGGAATCATTAACAAGAGTCCCTG 182

Query 244 ACATGGGCAGAAGGAGGTGGCCTGTGCATGCTGCTCCAGACCAAGGTGGCAG 303
Sbjct 183 ACATGGCC---AAGGAGGTGACCGGACTGCACCAGCTGTTCCAGACCAAGGTGGCAG 238

Query 304 CTCTACTGGGGGAGCCTGCTACTTCTGACAGGTATGGGGGGCATCCTGGGCCATTG 363


Sbjct 340 GTCTCAACTGTGGGGACCCTGTGTTTTTGAGAGTCATTGAGGACTGCGTCCTGGGTCTATTT 399
Query 864 TGGGTGTTCTTTTCAGGAGTTACTAGAGCCGCTTCACCCAGTGCAGGTATGTTTTGG 923
Sbjct 400 TGACCATCCTACTTCGGGACTGACGACGGCACTTTCCAATTGGCTCAGTCCGGGGG 459
Query 924 ACGGTTGGATGATGGTCTTGCAGAGGTGGAAGGGATGTGGGGGCTGCTGTGCTCCTCGC 983
Sbjct 460 ACTCTGCACTTTTGTTGTTGGAAAAAGAGGGAAGGTGGAGGTCTGTGTTCTCAG 519
Query 984 GCTCCAACCCAGGGTAGCCAGCTGAGTTGCAAGGAGATGGCTTTCTCAGGGACATTGA 1043
Sbjct 520 GCTCGAGCGCAAGGTAGCAGGTCTGGTCTGAAGAAGATGGGCTTCTCAAGGGTCTTGG 579
Query 1044 GCCACTCTGGAGCTGGAGTTGTAAGCCAGGGGCGCCAATGCACAGTCGGGTTTTTTGTGTG 1103
Sbjct 580 CCTATTAGAGGTTGGATGTGCGAACCCTCGGGCCTCAAGCCATCGGGCTTCTTCTGCG 639
Query 1104 TGGACGAGGGAGGCTGCCCCACACCCAGAGGTCTGTTGAGGTCATCTCCTTTCTGATT 1163
Sbjct 640 TGAACGAGGGCGGTCTTCCTTTTTTCAGCGGTTGCTGGATGTCATCTCTATATCTAACT 699
Query 1164 GCTCCAAAGCCTTTGTTGGCGCCATCTGCCAAGACTTTGCCGGAGGAAGGTGCCAG 1223
Sbjct 700 CTCCTAATAGCCGAATTCCTGACTGCCACCTGCCAAACTATGCCGAGGAACCTGCCGG 759
Query 1224 TGGACCGCATCGTGGAAGGCCTTAACAGCATTGGGCGGTTCGTGGTGGAATCAACC 1283
Sbjct 760 TGGATCTCATTGTGGGGGGCGAGAGACGCCTGGCGAAGATTGCCATGCAGGTCAGCC 819
Query 1284 TTCGATGAGTGGAGCACACCTCTGTTGGGGATCCTCCGGGACTCTCCGGGACTGCTTGA 1343
Sbjct 820 TCGTTACGATGGGGACTCACCTTGTGTGGGGATCCCTGTCTGTCCGGGACTGGTTCTGA 879
Query 1344 CAGCCGTCCACTGCTTCCCGGAGCGGAACCGGGTCCTGTCCCATGGCGAGTGTTCGCCG 1403
Sbjct 880 CCGCTGTCACGACGTGCTTTCCAGAGAGGAACCGGGTCCTATGTGGGTGGCGAGTATTCGCTG 939
Query 1404 GTGCCTGTCCCAGGCCTCTCCCCAGGAGTCTGCAGCTGGGGGGTGCAGGCTGTGTATTAGC 1463
Sbjct 940 GTGCTGTAGCCCGGACCTCACCTCATCCGTGCAGCTGGGGGTCCAGGCCTGTGATCTATC 999
Query 1464 ACTGGGGCTATCTTGCGTTCCGGAGACCCAACAGCAAGGTCAACAGCAAGGATATTGCCC 1523
Sbjct 1000 ATGGGGGCTACCTCCCTTCTGAGAACCCTAGTATCGACGAAACAGCTAATGACTATTACC 1059
Query 1524 TGGTTCACCTTTCAGTCCCTGCCCCTCACAGAAATACATCTAGCCTGTGTTGCTCCCAG 1583
Sbjct 1060 TGGTCAACCTCTCAGTCCCTTCCTTTCACAAAATACATCCAGCCGGTTTGTCCCTTG 1119
Query 1584 CTGCCGGCCAGCCCTGGTGGATGGCAAGGTCTGTACCGTGACGGGCTGGGGCAATGCC 1643
Sbjct 1120 GTGGCGAGCAGGGCCGTGGACGGCAAGGTCTGTATACTGTACCCGGCTGGGTACACAC 1179
Query 1644 AGTACTATGGCCAGCAAGTCTGTCCTATGGAAACCAGATCAAGCTCAAGCATGTTCTGTGGTG 1703
Sbjct 1180 AGTCGTATGGGCCAGCAAGTCTGTCCTACGAGAGGCCCGGTCCCCATCATAAGCAAGG 1239
Query 1704 ATGTCTGCAATGGCGCTGACTTCCATGGAAACCAGATCAAGCTCAAGCATGTTCTGTGGTG 1763
Sbjct 1240 ATGTCTGCAATGGCGCTGACTTCCATGGAAACCAGATCAAGCTCAAGCATGTTCTGTGGTG 1299

```
Sbjct  1140  AADTTTDDAACDADDDDGAATTCTACGGGAATCAAAPCAAADDAAAAATGTTCTTTCTG  1195

Query  1254  GTTACCCGAGGGGTGGCATTGATGCTTGTCAGGGTGAAGAAGGTGGTCCTTTGTGTGTG  1373
              |||| || |||||||| |||||||| ||||||| |||||||| ||| |||||| ||||
Sbjct  1200  GCTATCCTAAGGTTGGTATTGATGCATGCAGGGTAACAGCAGAAGGCACTTTTTATGTG  1259

Query  1314  AGAACAGGATTTTCCGGTGTCCACCTTGCCGACTCTTCTCCCATTTTGAGTTGGTGGTACTG  1433
              |||||| ||||| | ||| || | |||||| || |||||||  || ||||| || | |
Sbjct  1260  AGGAGAGAATCTCTGGAACATCAACATGGTGGCTCTGCTCATTGTAACCTGGGGTACTGG  1319

Query  1374  GCTGTGCCCTGTCCTAGAAGTCAGGGTTACACCAAAPTCAGTGACTTTCTGGAATTGAA  1493
              |||||||  |||||| |||||| || || ||||||||||||| ||||||||||||||||
Sbjct  1320  GCTGTGTTTTGGCCGGGAAGCCGGGAGTGTACACCAAAPCTATTGACTTCCGGGAGTGGA  1379

Query  1434  TCTTCGAGGCCATAAAGACTCACTCCGAAGCTACCGGCAPTGTAACTCAGCCCTGACC  1603
              ||||||||||||||||||||||||||||||| | |||||||| || |||| ||||||
Sbjct  1380  TCTTCCAGGCCATAAAGACTCACTCCGGAGCTACCGGCAPTGTAACTCAGCCCTGACC  1439
```

```
Score = 77.6 bits (40),  Expect = 1e-10
Identities = 175/230 (76%), Gaps = 9/230 (3%)
Strand=Plus/Plus

Query  1555  TCCACGCTGPPGCTGWGGATGGGACGCTTTTCTCTTTGGGCCCTGTTCACAGGTCCAAGGA  1614
              ||||| |||||||  |||| ||| |||||| || || |||| || ||||||||||||||
Sbjct  1317  TCTACACTCTGGCCTGGAAGGGAAGCGTTTTCTGCCTGPATCTGGTCGATAGATCGAAGGA  1376

Query  1615  CACCCTCCCTCCAGGGGTCCTCTTCCCACAGTGGCGGCCCACTCAAGCCCCAWGWCTAGC  1674
              | |  |  |||| || ||||||||||||| |||||||||  ||| | ||||
Sbjct  1377  TGC--TGGGTCCAAGGACCTTCTTCGCAGTGGCCGGCCACTCAATCCCAGTGGCCATT  1434

Query  1675  CAACCTCACCCTCCTGACCCCCATGTAAATATTGTTCTCCTCCTGGGA-CTCCTGTCTA  1733
              ||||||||||||  ||||||||||||| ||||  |||||  || || | |
Sbjct  1435  GG-CCTCACCCTCCCA---CCCCATCTAAATATTATCTCGTCCTCTGGGGGCTCTCTTTCTGA  1491

Query  1734  GGTGCCCTGATGATGGGATGCTCTTTAAATAATAAAGATGGTTTTGATT  1783
              || |||||  |:  ||||||||||||||||||||| ||||||||||
Sbjct  1492  GGCGCCCCTTGTGC--GGATGCTCTTTAAATAATAAAGATGGTTTTGATT  1539
```

```
Score = 66.1 bits (34),  Expect = 4e-07
Identities = 48/55 (87%), Gaps = 0/55 (0%)
Strand=Plus/Plus

Query  89    GGCCGCCCCACTGCTGCGGGGGCCACAGCTCCTCCCCACAGCCTGAAGATGACCC  143
             || ||||||||||| |||| |||||||| ||| ||||||| ||||||| ||||
Sbjct  45    GGCTGCCCGCTGCTGTGGGGACATCATCCCCTTCCCAGCCCGGAGACTRACCC  99
```

```
CPU time:     0.03 user secs.       0.00 sys. secs          0.03 total secs.

Lambda      K        H
    1.33   0.621     1.12

Gapped
Lambda      K        H
    1.33   0.621     1.12


Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 5, Extension: 2
Number of Sequences: 1
Number of Hits to DB: 147
Number of extensions: 9
Number of successful extensions: 6
Number of sequences better than 10.0: 1
Number of HSP's gapped: 3
Number of HSP's successfully gapped: 3
Length of query: 1788
Length of database: 18,244,403,544
Length adjustment: 27
Effective length of query: 1758
Effective length of database: 18,244,403,517
Effective search space: 32057173278282
Effective search space used: 32057173278282
X1: 11 (21.1 bits)
X2: 26 (50.0 bits)
X3: 26 (50.0 bits)
S1: 14 (27.6 bits)
S2: 21 (42.6 bits)
```